OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: November 7, 2003, 16:02:58 ; Search time 28 Seconds
(without alignments)
3412.068 Million cell updates/sec

Title: US-09-704-272-2
Perfect score: 11785
Sequence: 1 MACWPQLRLLLWKNLTFRRR..........VDVAVLTSFLQDEKVKESYV 2258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6901	58.6	1375	3	US-08-665-259-26	Sequence 26, Appl
2	6901	58.6	1375	3	US-09-634-238-243	Sequence 243, App
3	3112.5	26.4	1457	3	US-09-120-513-2	Sequence 2, Appli
4	3112.5	26.4	1457	3	US-09-450-105-2	Sequence 2, Appli
5	2618	22.2	1684	3	US-08-665-259-25	Sequence 25, Appl
6	2618	22.2	1684	3	US-08-762-500-27	Sequence 27, Appl
7	2618	22.2	1704	3	US-08-762-500-25	Sequence 25, Appl
8	359	3.0	315	4	US-08-762-500-75	Sequence 75, Appl
9	355	3.0	322	4	US-09-328-352-4388	Sequence 4388, App
10	353	3.0	322	4	US-09-107-532A-4662	Sequence 4662, Ap
11	352.5	3.0	335	4	US-09-107-532A-3752	Sequence 3752, Ap
12	344	2.9	589	4	US-09-252-991A-20837	Sequence 20837, Ap
13	338.5	2.9	1280	2	US-09-328-352-7592	Sequence 7592, Ap
14	338	2.9	402	4	US-09-583-276-19	Sequence 19, Appl
15	337.5	2.9	1280	4	US-09-107-532A-5360	Sequence 5360, Ap
16	337.5	2.9	1280	6	US-09-767-594-2	Sequence 2, Appli
17	335.5	2.8	291	4	US-5206352-4	Patent No. 5206352
18	335	2.8	1279	2	US-08-784-69A-2	Sequence 4205, App
19	335	2.8	1280	2	US-09-107-532A-4205	Sequence 4205, Appl
20	326.5	2.8	1280	2	US-09-316-167-2	Sequence 2, Appli
21	323	2.7	929	4	US-09-252-991A-22946	Sequence 22946, A
22	322.5	2.7	788	4	US-09-252-991A-28171	Sequence 28171, A
23	319.5	2.7	607	4	US-09-252-991A-18351	Sequence 18351, A
24	318	2.7	254	4	US-09-107-532A-4983	Sequence 4983, Ap
25	317	2.7	406	4	US-09-134-001C-3570	Sequence 3570, Ap
26	316.5	2.7	309	4	US-09-252-991A-21204	Sequence 21204, Ap
27	316.5	2.7	391	4	US-09-252-991A-20275	Sequence 20275, A
28	316.5	2.7	460	4	US-09-134-001C-3369	Sequence 3369, Ap
29	316	2.7	319	4	US-09-634-238-243	Sequence 243, App
30	310.5	2.6	1275	3	US-09-120-513-2	Sequence 2, Appli
31	310.5	2.6	1275	3	US-09-450-105-2	Sequence 2, Appli
32	310	2.6	1302	1	US-08-232-537-2	Sequence 2, Appli
33	302	2.6	329	4	US-09-107-532A-4844	Sequence 4844, Ap
34	299	2.5	350	4	US-09-107-532A-6978	Sequence 6978, Ap
35	299	2.5	528	4	US-09-107-532A-6340	Sequence 6340, Ap
36	296.5	2.5	304	4	US-09-107-532A-5424	Sequence 5424, Ap
37	293	2.5	337	4	US-09-134-001C-5550	Sequence 5550, Ap
38	287.5	2.4	1408	1	US-08-612-521-2	Sequence 2, Appli
39	287	2.4	1349	2	US-09-627-376-12	Sequence 12, Appl
40	285.5	2.4	233	4	US-09-134-001C-3824	Sequence 3824, Ap
41	284	2.4	231	4	US-09-134-001C-3292	Sequence 3292, Ap
42	281.5	2.4	288	4	US-09-107-532A-7040	Sequence 7040, Ap
43	280.5	2.4	307	4	US-09-134-001C-3539	Sequence 3539, Ap
44	280.5	2.3	257	4	US-09-134-001C-3539	Sequence 3539, Ap
45	276	2.3	303	4	US-09-107-532A-5192	Sequence 5192, Ap

RESULT 1
US-08-665-259-26
Sequence 26, Application US/08665259
Patent No. 6028173
GENERAL INFORMATION:
APPLICANT: Landes, Gregory M.
APPLICANT: Burn, Timothy C.
APPLICANT: Connors, Timothy D.
APPLICANT: Dackowski, William R.
APPLICANT: Van Raay, Terence J.
APPLICANT: Klinger, Katherine W.
TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,259
FILING DATE: 17-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 37,315
REFERENCE/DOCKET NUMBER: IG5-9.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1375 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: unknown
MOLECULE TYPE: protein

US-08-665-259-26

ALIGNMENTS

Query Match 58.6%; Score 6901; DB 3; Length 1375;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1332; Conservative 21; Mismatches 22; Indels 0; Gaps 0;

OM protein - protein search, using sw model

```
                        GenCore version 5.1.6
                        Copyright (c) 1993 - 2003 Compugen Ltd.

Run on:          November 7, 2003, 16:01:07 ; Search time 25 Seconds
                                             (without alignments)
                                             4247.452 Million cell updates/sec

Title:           US-09-704-272-2
Perfect score:   11785
Sequence:        1 MACWPQLRLLLWKNLTFRRR..........VDVAVLTSFLQDEKVKESYV 2258

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters:     127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_41:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.
```

[handwritten: Date]   [handwritten: "CERP"]

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description	
1	11718.5	99.4	2261	1	ABC1_HUMAN	O95477	homo sapien
2	11203.5	95.1	2261	1	ABC1_MOUSE	P41233	mus musculu
3	5836	49.5	2273	1	ABCR_HUMAN	P78363	homo sapien
4	4217	35.8	2436	1	ABC2_HUMAN	Q9BZC7	homo sapien
5	4073	34.6	2434	1	ABC2_MOUSE	P41234	mus musculu
6	2618	22.2	1704	1	ABC3_HUMAN	Q9Y758	homo sapien
7	1523	12.9	1704	1	CED7_CAEEL	Q03434	caenorhabdi
8	411	3.5	330	1	DRRA_STREP	P32010	streptomyce
9	380.5	3.2	343	1	NDI1_RHISN	P55476	rhizobium s
10	366	3.1	304	1	NDI1_RHIS3	P72335	rhizobium s
11	347	2.9	308	1	YADG_ECOLI	P36879	escherichia
12	347	2.9	335	1	NDI2_RHIME	Q8GRH6	rhizobium s
13	345	2.9	340	1	NDI1_RHILO	P23703	rhizobium m
14	343.5	2.9	347	1	NDI1_RHIGA	P50032	rhizobium g
15	337.5	2.9	1280	1	MDRI_HUMAN	P08183	homo sapien
16	336	2.9	894	1	YHIH_ECOLI	P37624	escherichia
17	331.5	2.9	355	1	NDII_RHIME	O52618	rhizobium m
18	331.5	2.8	578	1	YBHF_ECOLI	P75776	escherichia
19	329.5	2.8	308	1	NOSF_PSEST	P19844	pseudomonas
20	327.5	2.8	1281	1	MDR3_CRIGR	P23174	cricetulus
21	327.5	2.8	306	1	NODI_BRAJA	P26050	bradyrhizob
22	324.5	2.8	1276	1	MDR2_MOUSE	P21440	mus musculu
23	324.5	2.8	1276	1	MDR1_CRIGR	P21448	cricetulus
24	323	2.7	1362	1	PMDI_SCHPO	P36619	schizosach
25	321.5	2.7	1276	1	MDR3_MOUSE	P21447	mus musculu
26	320	2.7	1276	1	MDR2_MOUSE	P21449	mus musculu
27	317.5	2.7	1276	1	MDR1_CRIGR	P08720	cricetulus
28	317.5	2.7	311	1	NODI_RHILV	P22040	rhizobium l
29	315.5	2.7	354	1	Y415_SYNY3	P43245	synechocyst
30	313	2.7	1277	1	MDR1_RAT	Q08201	rattus norv
31	313	2.7	1278	1	MDR2_RAT	P06795	rattus norv
32	309	2.6	1276	1	MDR1_MOUSE	Q9Z313	mus musculu
33	304.5	2.6	1294	1	YOH5_YEAST	Q08234	saccharomyc
34	301	2.6	381	1	OPBA_BACSU	Q45460	bacillus su
35	299	2.5	380	1	OPCA_BACSU	O34992	bacillus su
36	296	2.5	335	1	Y719_ANASP	Q05067	anabaena sp
37	292.5	2.5	262	1	YA23_METJA	Q58429	methanococc
38	292.5	2.5	305	1	YHCH_BACSU	P54592	bacillus su
39	290.5	2.5	274	1	Y179_MYCPN	Q50294	mycoplasma
40	288	2.4	1321	1	AB11_HUMAN	O95342	homo sapien
41	287.5	2.4	274	1	Y179_MYCGE	P47425	mycoplasma
42	287.5	2.4	1279	1	MDR3_HUMAN	P21439	homo sapien
43	284.5	2.4	1279	1	MDR1_CAEEL	P34712	caenorhabdi
44	284	2.4	306	1	BCRA_BACLI	P42232	bacillus li
45	282.5	2.4	1336	1	MAM1_SCHPO	P78966	schizosach

RESULT 1
ABC1_HUMAN

```
ID   ABC1_HUMAN    STANDARD;    PRT;   2261 AA.
AC   O95477; Q96S56; Q96T85; Q9NQV4; Q9UN06; Q9UN07; Q9UN08; Q9UN09;
AC   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   ATP-binding cassette, sub-family A, member 1 (ATP-binding cassette
DE   transporter 1) (ATP-binding cassette 1) (ABC-1) (Cholesterol efflux
DE   regulatory protein).
GN   ABCA1 OR ABC1 OR CERP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20345099; PubMed=10884428;
RA   Santamarina-Fojo S., Peterson K.M., Knapper C.L., Qiu Y.,
RA   Freeman L.A., Cheng J.-F., Osorio J., Remaley A.T., Yang X.-P.,
RA   Haudenschild C.C., Prades C., Chimini G., Blackmon E.B.,
RA   Francois T.L., Duverger N., Rubin E.M., Rosier M., Denefle P.,
RA   Fredrickson D.S., Brewer H.B. Jr.;
RT   "Complete genomic sequence of the human ABCA1 gene: analysis of the
RT   human and mouse ATP-binding cassette A promoter.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:7987-7992(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RA   Schwartz K., Lawn R.M., Wade D.P.;
RT   "ABCA1 gene expression and apoA-I-mediated cholesterol efflux are
RT   regulated by LXR.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21251004; PubMed=11352567;
RA   Tanaka A.R., Abe-Dohmae S., Arakawa R., Sadanami K., Kidera A.,
RA   Kioka N., Amachi T., Yokoyama S., Ueda K.;
RT   "A new topological model of functional human ABCA1-signal peptide
RT   cleavage and glycosylation of a large extracellular domain.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Qiu Y., Cavelier L., Chiu S., Yang X., Rubin E., Cheng J.-F.;
RT   "Human and mouse ABCA1 comparative sequencing and transgenesis
RT   studies revealing novel regulatory sequences.";
RL   Genomics 73:66-76(2001).
RN   [5]
RP   SEQUENCE OF 21-2261 FROM N.A.
RX   MEDLINE=99194549; PubMed=10092505;
RA   Langmann T., Klucken J., Reil M., Liebisch G., Luciani M.F.,
RA   Chimini G., Kaminski W.E., Schmitz G.;
RT   "Molecular cloning of the human ATP-binding cassette transporter 1
RT   (hABC1): evidence for sterol-dependent regulation in macrophages.";
RL   Biochem. Biophys. Res. Commun. 257:29-33(1999).
RN   [6]
```

ALIGNMENTS

OM protein - protein search, using sw model

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.6

Run on: November 7, 2003, 16:02:12 ; Search time 35 Seconds
(without alignments)
6204.254 Million cell updates/sec

Title: US-09-704-272-2
Perfect score: 11785
Sequence: 1 MACWPQLRLLLWKNLTFRRR..........VDVAVLTSFLQDEKVKESYV 2258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	10869.5	92.2	2201	2	A54774	ATP binding casset
2	3332.5	28.3	1529	2	A59189	ATP-binding casset
3	3112.5	26.4	1472	2	B54774	ATP-binding casset
4	2621	22.2	1704	2	S71363	probable ATP-bindi
5	2618	22.2	1704	2	A59188	ATP-binding casset
6	2043.5	17.3	1802	2	T33783	probable ATP-bindi
7	1985.5	16.8	1816	2	A84845	hypothetical prote
8	1858	15.8	1447	2	T15200	probable ABC trans
9	1778.5	15.1	1317	2	C88925	hypothetical prote
10	1530	13.0	1704	2	F88559	protein F33E11.4 [
11	1523	13.0	1758	2	T42749	protein C48B4.4b [
12	1520.5	12.9	1767	2	S60124	transport protein
13	1398	11.9	1246	2	T00826	hypothetical prote
14	1153	9.8	1564	2	T21121	hypothetical prote
15	1017	8.6	1431	2	T22748	hypothetical prote
16	1013.5	8.6	373	2	T41150	hypothetical prote
17	846	7.2	269	2	T46467	hypothetical prote
18	784	6.7	1011	2	T07712	hypothetical prote
19	684.5	5.8	900	2	T07717	probable ABC-type
20	662.5	5.6	895	2	T07714	probable ABC-type
21	649	5.5	722	2	T07716	probable ABC-type
22	638.5	5.4	925	2	T07713	probable ABC-type
23	619	5.3	1336	2	T18288	ABC transport prot
24	485	4.1	196	2	T12512	hypothetical prote
25	432	3.7	339	2	S74048	probable daunorubi
26	426.5	3.6	324	2	C71081	probable resistanc
27	422.5	3.6	664	2	T07715	probable ABC-type
28	419.5	3.6	328	2	E75108	daunorubicin resis
29	412.5	3.5	310	2	C96929	ABC-type multidrug
30	411	3.5	330	2	S27707	daunorubicin resis
31	410.5	3.5	327	2	D72257	hypothetical prote
32	410	3.5	314	2	D97318	ABC-type MDR trans
33	409.5	3.5	398	2	C69485	daunorubicin resis
34	404.5	3.4	297	2	AE1816	ABC transporter (A
35	399.5	3.4	333	2	D72492	probable ABC trans
36	395.5	3.4	310	2	E96920	ABC transporter (A
37	395	3.4	311	2	G69803	ABC transporter (A
38	394	3.3	246	2	S75436	hypothetical prote
39	390	3.3	301	2	E72384	ABC transporter, A
40	388	3.3	331	2	D70984	probable drrA prot
41	387	3.3	259	2	H97301	ABC-type MDR trans
42	386	3.3	350	2	S73019	daunorubicin resis
43	386	3.3	331	2	B69065	ABC transporter (A
44	380	3.2	312	2	H97342	ABC-type MDR trans
45	378	3.2	130	2	I38906	ATP-binding casset

ALIGNMENTS

RESULT 1
A54774
ATP binding cassette transporter ABC1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 05-Apr-1995 #sequence_revision 05-Apr-1995 #text_change 02-Feb-2001
C;Accession: A54774
R;Luciani, M.F.; Denizot, F.; Savary, S.; Mattei, M.G.; Chimini, G.
Genomics 21, 150-159, 1994
A;Title: Cloning of two novel ABC transporters mapping on human chromosome 9.
A;Reference number: A54774; MUID:94375008; PMID:8088782
A;Accession: A54774
A;Molecule type: mRNA
A;Residues: 1-2201 <LUC>
A;Cross-references: GB:X75926; NID:g495556; PIDN:CAA53530.1; PID:g495257
A;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolo
C;Keywords: ATP; duplication; nucleotide binding; P-loop
F;856-1047/Domain: ATP-binding cassette homology <ABC1>
F;873-880/Region: nucleotide-binding motif A (P-loop)
F;1869-2060/Domain: ATP-binding cassette homology <ABC2>
F;1886-1893/Region: nucleotide-binding motif A (P-loop)

Query Match 92.2%; Score 10869.5; DB 2; Length 2201;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 2084; Conservative 54; Mismatches 60; Indels 3; Gaps 1;

```
Qy   61 MPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFENKSIVARLFSDARLLLYSQKDT 120
            ||||||||||||||||||||||||||||||||||||||||| |||||| |||||| ||
Db    1 MPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFENKSIVSRLFSDAQRLLYSQRDT 60

Qy  121 SMKDMRKVLRTLQQIKKSSSNLKLQDFLVDNETFSGFLYHNLSLPKSTVDKMLRADVILH 180
            |:||| ||||:|:||| .||||||||||||||||| :|||||| ||||. ||..:|:
Db   61 SIKDMHKVLRMLRQIKHPNSNLKLQDFLVDNETFSGEIQHNLSLPRSTVDSLLQXNVGLQ 120

Qy  181 KVFLQGYQLHITSLCNGSKSEEMIQLGDQEVSELCGLPREKLAAAERVLRSNMDILKPIL 240
            ||||||||||::||||||| ||:|||||.|||:||||||:||.|||||||.:|||||::
Db  121 KVFLQGYQLHASLCNGSKLEEIIQLGDAEVSALCGLPRKKLDAAERVLRYNMDILKPVV 180

Qy  241 RTLNSTSPPSKELAEATKTLLHSLGTLAQELFSMRSWSDMRQEVMFLTNVNSSSSSTQI 300
            .||||||.| :|||||..|||.||| |||||||:.||||||||||||||||||||||||
Db  181 TKLNSTSHLPTQHLAEATTVLLDSLGGLAQELFSTKSWSDMRQEVMFLTNVNSSSSSTQI 240

Qy  301 YQAVSRIVCGHPEGGGLKIKSLNWYEDNNYKALFGGNGTEEDAEFFYDNSTTPYCNDLMK 360
            |||||||||.||||||||||||||||||||||||||:|||||.: ||:|||||||||||
Db  241 YQAVSRIVCCHPEGGGLKIKSLNWYEDNNYKALFGGNNTEEDVDTFYPNSTTPYCNDLMK 300

Qy  361 NLESSPLSRIIWKALKPLLVGKILYTPDTPATRQVVAEVNKTFQELAVFHDLEGMWEELS 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 NLESSPLSRIIWKALKPLLVGKILYTPDTPATRQVVAEVNKTFQELAVFHDLEGMWEELS 360

Qy  421 PKIWTFMENSQEMDLVRMLLDSRDNDHFWEQQLDGLLDWTAQDIVAFLAKHPEDVQSSNGS 480
            |:||||||||||||||| ||||| || ||||:|||||||||||:.||||::||||||.|||
Db  361 PQIWTFMENSQEMDLVRTLLDSRGNDQFWEQKLDGLLDWTAQDIMAFLAKNPEDVQSPNGS 420
```

OM protein - protein search, using sw model

Copyright (c) 1993 - 2003 Compugen Ltd.
GenCore version 5.1.6

Run on: November 7, 2003, 16:04:18 ; Search time 48 Seconds
(without alignments)
8079.373 Million cell updates/sec

Title: US-09-704-272-2
Perfect score: 11785
Sequence: 1 MACWPQLRLLLWKNLTFRRR...........VDVAVLTSFLQDEKVKESYV 2258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	11734.5	99.6	2261	12	US-10-313-641-9	Sequence 9, Appli
2	11734.5	99.6	2261	12	US-10-313-641-10	Sequence 10, Appl
3	11715.5	99.4	2261	10	US-09-995-542-11	Sequence 11, Appl
4	11715.5	99.4	2261	10	US-09-846-456-11	Sequence 11, Appl
5	11715.5	99.4	2261	10	US-09-984-827-2	Sequence 2, Appli
6	11714.5	99.4	2261	11	US-09-984-827-129	Sequence 129, App
7	11714.5	99.4	2261	11	US-09-984-827-136	Sequence 136, App
8	11712.5	99.4	2261	11	US-09-984-827-127	Sequence 127, App
9	11712.5	99.4	2261	11	US-09-984-827-128	Sequence 128, App
10	11712.5	99.4	2261	11	US-09-984-827-131	Sequence 131, App
11	11712.5	99.4	2261	11	US-09-984-827-134	Sequence 134, App
12	11712.5	99.4	2261	11	US-09-984-827-135	Sequence 135, App
13	11710.5	99.3	2261	11	US-09-984-827-133	Sequence 133, App
14	11705.5	99.3	2261	11	US-09-984-827-130	Sequence 130, App
15	11704.5	99.3	2261	11	US-09-984-827-132	Sequence 132, App
16	11176.5	94.8	2261	12	US-10-340-097-118	Sequence 118, App
17	11176.5	94.8	2261	12	US-10-336-215-118	Sequence 118, App
18	11176.5	94.8	2261	12	US-10-336-219-118	Sequence 118, App
19	10869.5	92.2	2201	10	US-09-995-542-9	Sequence 9, Appli
20	5838.5	49.5	2310	10	US-09-995-542-10	Sequence 10, Appl
21	5820	49.4	2273	10	US-09-995-542-12	Sequence 12, Appl
22	5799	49.2	2273	12	US-10-340-097-3	Sequence 3, Appli
23	5799	49.2	2273	12	US-10-336-215-3	Sequence 3, Appli
24	5799	49.2	2273	12	US-10-336-219-3	Sequence 3, Appli
25	5716	48.5	2146	9	US-09-995-542-5	Sequence 5, Appli
26	5703	48.4	2144	9	US-09-858-194-2	Sequence 2, Appli
27	5703	48.4	2144	12	US-10-154-419-2	Sequence 2, Appli
28	5702	48.4	2235	12	US-10-340-097-6	Sequence 6, Appli
29	5702	48.4	2235	12	US-10-336-215-6	Sequence 6, Appli
30	5702	48.4	2235	12	US-10-336-219-6	Sequence 6, Appli
31	5680.5	48.2	2167	10	US-09-995-542-2	Sequence 2, Appli
32	5558	47.2	2100	10	US-09-995-542-4	Sequence 4, Appli
33	5530.5	46.9	2121	10	US-09-995-542-3	Sequence 3, Appli
34	4227	35.9	2436	9	US-09-799-693-8	Sequence 8, Appli
35	4227	35.9	2436	15	US-10-156-239-8	Sequence 8, Appli
36	4227	35.9	2436	15	US-10-199-485-8	Sequence 8, Appli
37	4014.5	34.1	2001	14	US-10-072-621-8	Sequence 8, Appli
38	3779.5	32.1	1550	10	US-09-995-542-8	Sequence 8, Appli
39	3235.5	27.5	2595	15	US-10-072-900-5	Sequence 5, Appli
40	3110.5	26.4	1472	12	US-10-340-097-119	Sequence 119, App
41	3110.5	26.4	1472	12	US-10-336-215-119	Sequence 119, App
42	3110.5	26.4	1472	12	US-10-336-219-119	Sequence 119, App
43	3054	25.9	2516	12	US-10-340-900-6	Sequence 6, Appli
44	2625	22.3	1704	12	US-10-340-097-120	Sequence 120, App
45	2625	22.3	1704	12	US-10-336-215-120	Sequence 120, App

ALIGNMENTS

RESULT 1
US-10-313-641-9

; Sequence 9, Application US/10313641
; Publication No. US20030162758A1
; GENERAL INFORMATION:
; APPLICANT: Ishida, Brian
; APPLICANT: Duncan, Keith
; APPLICANT: Bailey, Kathy
; APPLICANT: Kane, John
; APPLICANT: Schwartz, Daniel
; TITLE OF INVENTION: Treatments for Age Related-Macular Degeneration (AMD)
; FILE REFERENCE: PO235IUS2
; CURRENT APPLICATION NUMBER: US/10/313,641
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/415,864
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US 60/340,498
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 9
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Human
US-10-313-641-9

Query Match 99.6%; Score 11734.5; DB 12; Length 2261;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2253; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

```
Qy    1 MACWPQLRLLLWKNLTFRRRQTCQLLEVAWPLIFLIIISVRLSYPPYEQHECHFPNKA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MACWPQLRLLLWKNLTFRRRQTCQLLEVAWPLIFLIIISVRLSYPPYEQHECHFPNKA 60

Qy   61 MPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIVARLFSDARLLLYSQKDT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 MPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIVARLFSDARLLLYSQKDT 120
```

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: November 7, 2003, 16:01:37 ; Search time 64 Seconds
                     9104.418 Million cell updates/sec
                     (without alignments)

Title: US-09-704-272-2
Perfect score: 11785
Sequence: 1 MACWPQLRLLLWKNLTFRRR..........VDVAVLTSFLQDEKVKESYV 2258

Scoring table: BLOSUM62
                     Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SPTREMBL_23:*
            1: sp_archea:*
            2: sp_bacteria:*
            3: sp_fungi:*
            4: sp_human:*
            5: sp_invertebrate:*
            6: sp_mammal:*
            7: sp_mhc:*
            8: sp_organelle:*
            9: sp_phage:*
           10: sp_plant:*
           11: sp_rodent:*
           12: sp_virus:*
           13: sp_vertebrate:*
           14: sp_unclassified:*
           15: sp_rvirus:*
           16: sp_bacteriap:*
           17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	10165	86.3	2260	13	Q8UVV4	Q8uvv4 gallus gall
2	7334.5	62.2	1487	11	Q8BPY1	Q8bpy1 mus musculu
3	5838.5	49.5	2019	11	O35600	O35600 mus musculu
4	5825	49.4	2281	6	O02698	O02698 bos taurus
5	5723	48.6	2146	4	Q9B2C4	Q9b2c4 homo sapien
6	5711	48.5	2146	4	Q9NR73	Q9nr73 homo sapien
7	5706	48.4	2146	4	Q8IZY2	Q8izy2 homo sapien
8	5681.5	48.2	2159	11	Q91V24	Q91v24 mus musculu
9	5323.5	45.2	2008	4	Q96S58	Q96s58 homo sapien
10	4232	35.9	2434	11	Q9ESR9	Q9esr9 rattus norv
11	4227	35.9	2436	11	Q9HC28	Q9hc28 homo sapien
12	3209	27.2	2347	4	Q8IZW6	Q8izw6 homo sapien
13	3200	27.2	2277	4	Q96JT3	Q96jt3 homo sapien
14	2698.5	22.9	1764	5	Q8MUA3	Q8mua3 strongyloce
15	2562.5	21.7	1704	11	Q8R420	Q8r420 mus musculu
16	2463	20.9	1713	5	Q9VRG4	Q9vrg4 drosophila
17	2146.5	18.2	1846	10	Q8W010	Q8w010 arabidopsis
18	2043.5	17.3	1802	5	Q9TXV8	Q9txv8 caenorhabdi
19	2024	17.2	1547	5	O01790	O01790 caenorhabdi
20	2019	17.1	1660	5	Q9VVJ9	Q9vvj9 drosophila
21	2015.5	17.1	1843	5	Q9GQS2	Q9gqs2 leishmania
22	2007.5	17.0	867	4	Q96HC2	Q96hc2 homo sapien
23	1984	16.8	1896	5	Q8WS98	Q8ws98 leishmania
24	1918	16.3	1621	5	Q8T6J5	Q8t6j5 dictyosteli
25	1828	15.5	1631	5	Q8T6J1	Q8t6j1 dictyosteli
26	1819.5	15.4	1608	5	Q8T6J2	Q8t6j2 dictyosteli
27	1807	15.3	1662	5	Q8T6J3	Q8t6j3 dictyosteli
28	1800	15.3	1750	5	Q9BKL1	Q9bkl1 trypanosoma
29	1794.5	15.2	1702	5	Q8T6J4	Q8t6j4 dictyosteli
30	1527.5	13.0	1526	5	Q9VVK6	Q9vvk6 drosophila
31	1502.5	12.7	1642	11	Q9NY14	Q9ny14 homo sapien
32	1469	12.5	1642	4	Q8WW27	Q8ww27 homo sapien
33	1467	12.4	1642	11	Q8K448	Q8k448 mus musculu
34	1443	12.2	1642	11	Q8S8T6	Q8s8t6 arabidopsis
35	1413	12.0	1222	10	Q8K449	Q8k449 mus musculu
36	1411.5	12.0	1623	11	Q8WW24	Q8ww24 homo sapien
37	1408	11.9	1543	4	Q8IUA7	Q8iua7 homo sapien
38	1369.5	11.6	1624	4	Q8N139	Q8n139 homo sapien
39	1367.5	11.6	1617	4	Q8IUA7	Q8iua7 homo sapien
40	1364.5	11.6	1624	4	Q8WWZ5	Q8wwz5 homo sapien
41	1358.5	11.5	1617	4	Q8WW26	Q8ww26 homo sapien
42	1349.5	11.5	1620	11	Q8K440	Q8k440 mus musculu
43	1334.5	11.3	1620	11	Q8K442	Q8k442 mus musculu
44	1327.5	11.3	1581	4	O94911	O94911 homo sapien
45	1320	11.2	1624	11	Q8K441	Q8k441 mus musculu

ALIGNMENTS

RESULT 1
Q8UVV4

ID   Q8UVV4   PRELIMINARY;   PRT;   2260 AA.
AC   Q8UVV4;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   ATP-binding cassette transporter 1.
GN   ABCA1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Attie A.D., Brooks-Wilson A.R., Walker D., McManus B.,
RA   Gray-Kellar M.P., Macdonald M.L.E., Roomp K., Tebon A., Zhang L.-H.,
RA   Mulligan J., Sensen C., Bitgood J.J., Cook M.E., Kastelein J.J.P.,
RA   Hayden M.R.;
RT   "Cholesterol Ester Accumulation in Hepatocytes and Intestinal Lamina
RT   Propria Caused by an ABCA1 Mutation in WHAM Chickens.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF362377; AAL56247.1; -.
DR   InterPro; IPR003593; AAA ATPase.
DR   InterPro; IPR003439; ABC_transporter.
DR   Pfam; PF00005; ABC_tran; 2.
DR   ProDom; PD000006; ABC_transporter; 2.
DR   SMART; SM00382; AAA; 2.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding.
SQ   SEQUENCE   2260 AA;   254070 MW;   19D137F342F98662 CRC64;

Query Match   86.3%;   Score 10165;   DB 13;   Length 2260;
Best Local Similarity   85.1%;   Pred. No. 0;
Matches 1929;   Conservative 163;   Mismatches 159;   Indels 16;   Gaps 8;

Qy   1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60

OM protein - protein search, using sw model

Copyright (c) 1993 - 2003 Compugen Ltd.
GenCore version 5.1.6

Run on: November 7, 2003, 16:00:12 ; Search time 59 Seconds
                                            (without alignments)
                                            6074.654 Million cell updates/sec

Title: US-09-704-272-2
Perfect score: 11785
Sequence: 1 MACWPQLRLLLWKNLLFRRR..........VDVAVLTSFLQDEKVKESYV 2258
Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 15872673 residues
Total number of hits satisfying chosen parameters: 1107863

Post-processing:
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   A_Geneseq_19Jun03:*

            1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
            2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
            3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
            4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
            5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
            6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
            7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
            8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
            9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
           10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
           11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
           12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
           13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
           14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
           15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
           16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
           17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
           18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
           19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
           20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
           21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
           22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
           23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
           24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Match	Query Length	DB Match	ID	Description
1	11737.5	99.6	2261	21	AAB38082	Human ABC1 cholest
2	11737.5	99.6	2261	22	AAB38113	Human ABC1 cholest
3	11737.5	99.6	2261	24	AAB71749	Human ABC1 protein
4	11734.5	99.6	2261	22	ABU01899	Human ATP-binding
5	11734.5	99.6	2261	21	AAB38109	Human ABC1 cholest
6	11734.5	99.6	2261	21	AAB38111	Human ABC1 cholest
7	11734.5	99.6	2261	21	AAB38114	Human ABC1 cholest
8	11734.5	99.6	2261	21	AAB38117	Human ABC1 cholest
9	11733.5	99.6	2261	21	AAB38105	Human ABC1 cholest
10	11733.5	99.6	2261	21	AAB38110	Human ABC1 cholest
11	11732.5	99.6	2261	21	AAB38113	Human ABC1 cholest
12	11732.5	99.5	2261	21	AAB38116	Human ABC1 cholest
13	11731.5	99.5	2261	21	AAB38112	Human ABC1 cholest
14	11730.5	99.5	2261	21	AAB31361	Human ABC1 cholest
15	11730.5	99.5	2261	22	AAB31365	Amino acid sequenc
16	11726.5	99.5	2261	22	AAB31362	Amino acid sequenc
17	11726.5	99.5	2261	21	AAB31366	Amino acid sequenc
18	11725.5	99.5	2261	21	AAB38104	Human ABC1 cholest
19	11723	99.5	2260	21	AAB38106	Human ABC1 cholest
20	11722.5	99.5	2261	22	AAB31363	Amino acid sequenc
21	11722.5	99.5	2259	21	AAB31367	Amino acid sequenc
22	11715.5	99.4	2261	22	AAB38107	Human ABC1 cholest
23	11715.5	99.4	2261	22	AAE13022	Human ABC1 FHA-3 m
24	11715.5	99.4	2261	22	AAM50228	Human ATP binding
25	11715.5	99.4	2261	22	AAU02176	Homo
26	11715.5	99.4	2261	22	AAU02177	Homo
27	11715.5	99.4	2261	23	AAB38311	Human ABC1 mutant
28	11715.5	99.4	2261	23	ABP52092	Homo sapiens ABC t
29	11715.5	99.4	2261	23	AAE22000	Human ABC1 full-le
30	11714.5	99.4	2261	23	ABB83117	Polymorphic human
31	11714.5	99.4	2261	23	ABB83124	Polymorphic human
32	11712.5	99.4	2261	23	AAU02183	Human ABC1 mutant
33	11712.5	99.4	2261	22	AAU02188	Human ABC1 mutant
34	11712.5	99.4	2261	23	ABB83115	Polymorphic human
35	11712.5	99.4	2261	23	ABB83116	Polymorphic human
36	11712.5	99.4	2263	23	ABB83119	Polymorphic human
37	11712.5	99.4	2261	23	ABB83122	Polymorphic human
38	11712.5	99.4	2261	23	ABB83123	Polymorphic human
39	11711.5	99.4	2261	22	AAU02189	Human ABC1 mutant
40	11710.5	99.4	2261	23	ABB83121	Polymorphic human
41	11709.5	99.4	2261	22	AAU02181	Human ABC1 mutant
42	11708.5	99.4	2261	22	AAM78550	Human protein SEQ
43	11708.5	99.4	2263	22	ABB11956	Polymorphic human
44	11708.5	99.4	2261	22	AAM79534	Human ABCA1 homolo
45	11707.5	99.3	2261	22	AAU02182	Human protein SEQ

ALIGNMENTS

RESULT 1
AAB38082

```
ID   AAB38082 standard; Protein; 2261 AA.
XX
AC   AAB38082;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 cholesterol transporter.
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal.
OS   Homo sapiens.
XX
PN   WO200005318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999; 99US-0124702.
PR   08-JUN-1999; 99US-0138048.
PR   17-JUN-1999; 99US-0139600.
PR   01-SEP-1999; 99US-0151977.
```